UNIVERSITY OF YORK

Department of Economics and Related Studies

HORIZONTAL INEQUITY IN THE DIRECT TAX SYSTEM: MEASUREMENT AND NORMATIVE ISSUES

Ana Sofia Ferreira[1]

MSc Dissertation in Economic and Social Policy Analysis

Supervisor: Professor Peter Lambert

September 1997

[1] The author benefits from a scholarship from the program PRAXIS XXI (Sub-Program Science and Technology): BM/8056/96.

0) INTRODUCTION

Taxation of personal income is typically seen as one of the most adequate forms of taxation, in terms of fairness criteria.

Progressivity in taxation, in modern tax systems , is essentially attained through the income tax, having two essential equity principles that will be developed and explored throughout this work: horizontal equity and vertical equity (leading to progressivity). In spite of the central role of both, we have chosen to place a greater emphasis on the analysis of the former.

Generally, vertical equity recommends differentiation of tax liabilities according to differences of tax unit´s (equivalent) income, while horizontal equity demands equality in the fiscal treatment of tax units in the same circumstances. Therefore, these are principles of distributional equity.

Despite the wide agreement, among politicians, academicians, and citizens in general, on the validity of these major principles, their practical implementation in tax systems is very often embodied in controversy since they are derived from normative views on the distributional aspects of taxation, the extent of which depends, ultimately, on subjective considerations about social justice, i.e. on personal tastes.

The identification of these normative views, and the analysis of the measurement issues related to the degree of existing vertical equity (linked to the progressivity of the tax system) and of horizontal inequity induced by the income tax will be the central concern of this work.

After having properly defined the idea of progressivity of the income tax, primary importance will be given to the concept of horizontal inequity, as well as to possible ways of assessing its magnitude in a certain tax system: measurement techniques -also embodying normative views and serious methodological problems- will be analysed and several measures of horizontal inequity will be presented. Finally, an empirical application will illustrate these latest aspects.

1) THE PROGRESSIVE INCOME TAX

The income tax is said to be progressive when the average rate of taxation that an income unit faces increases with its level of (pre-tax) income.

In mathematical terms, if we define $t(x)$ as the tax payment of an income unit with a level of income x, progressivity of $t(x)$ means that $t(x)/x$ is an increasing function of x.

Moreover, assuming differentiability of $t(x)$, progressivity of the income tax (in a strict sense) thus implies:

$$d\,[\,t(x)/x\,]\,/\,d\,x > 0\,,\ \text{for all } x.^{2}$$

By allowing $t(x) < 0$, we include here the possibility of negative income taxes, i.e. income-related benefits.

Analysis of progressivity is typically carried with the help of two strong assumptions, that do not hold in real world situations; on the one hand, that there is a common schedule $t(x)$ applying equally to all tax units[3], and on the other, that the population of income units is homogeneous. This is to say that we assume a population of comparable units, allowing (artificially) for abstraction from the issue of horizontal equity, in order to focus on vertical equity considerations. Even if heterogeinity of income units is allowed, progressivity analysis is then performed using equivalent income, with the further assumption that this coincides totally with all tax code requirements. In other words, distributional analysis of the income tax typically has implied a divorce between the concepts of vertical and horizontal equity. Progressivity measures (some of which will be mentioned below) are thus, in most of the literature on public finance, essentially linked to vertical equity, implicitly assuming the existence of perfect horizontal equity in the tax system under analysis, and therefore, leading

[2] If the income tax is proportional, the average tax rate is constant for any level of income and this implies an equality in this mathematical expression.

[3] We will use the terms "income unit", "tax unit" or "fiscal unit" as having exactly the same meaning: they refer to the unit -individual, family, household, equivalent adult- that pays the income tax. This aspect will be considered more deeply in section 3.2 of this work.

to biased conclusions in distributional assessment of income taxation.

Theoretically, the implications of a progressive income tax, in terms of it being inequality reducing and (comparatively) welfare improving, have been widely established and mostly verified in empirical studies.

Jakobsson and Fellman have separatelly proved that, under standard assumptions about the income tax, its progressivity means that it reduces inequality from the pre to the post-tax distribution of income.[4]

Assuming a homogeneous population (or subpopulation) of tax units, to which the tax code applies equally, the Jakobsson-Fellman theorem[5] shows that a progressive income tax implies disproportionality in tax payments among income units with distinct levels of pre-tax income, causing thus a greater degree of equality in the post-tax distribution of income. These two aspects of the tax -departure from proportionality and redistributive effect- are the basic features of its progressivity. Redistributive effect can be seen as a function of the level (related to the average tax rate) and of the pattern of taxes (linked to its departure from proportionality).This result has been shown by Kakwani (1976).

Considering now the impact on social welfare, this is typically assessed abstracting from the benefits that tax units receive from government´s role in modern economies, related namely with the production and provision of public or merit goods. In this context, (positive) income taxation can only reduce welfare: in spite of reducing inequality, it provokes a reduction of disposable income[6] for all income units. This conclusion is valid according to all Paretian social welfare functions (hereafter, SWF).

[4] See Jakobsson (1976) and/or Fellman (1976) for further details.
[5] For a mathematical proof of this result, see Lambert (1993), p.150.
[6] In this very partial equilibrium analysis, we stress again.

Consequently, having a generalised utilitarian SWF as the criterion for welfare analysis, progressivity is only seen as welfare improving by comparison with another way of taxation: a proportional income tax, raising exactly the same revenue (equal-yield proportional tax). After applying a progressive tax to some pre-tax distribution of income, the post tax income distribution will have the same mean after the imposition of an equal-yield proportional tax on the same pre-tax distribution; the progressive tax of course has a more equalizing effect, while the proportional tax is neutral (inequality-preserving). Using the well-known Atkinson´s theorem[7], we can conclude that a progressive income tax is comparatively welfare-improving[8], i.e. reduces welfare by less than a proportional tax raising the same revenue, from the same pre-tax distribution of income.

Progressive taxation is thus socio-economically seen as a way of reducing disparities of "well-being" among a population of income units, and it is intuitively accepted as an equitable method of public intervention. The polemic issue is, however, agreement among politicians, theorists and voters, on the "adequate" degree of progression of the income tax, which varies in time, place and also according to individuals´preferences and moral values.

The redistributive effect of a progressive income tax is identified with its vertical effect, when it is assumed that the tax causes no horizontal inequities (or reranking of income units). Progression is thus normally assessed having income redistribution theory as the underlying adequate framework for analysis.

Jakobsson (1976) quotes Musgrave and Thin (1948) and the four local measures of progression that they mention: average rate progression, marginal rate progression, liability progression and residual income progression. Notwithstanding, this author considers that any measure of progression is adequate only if progressivity of a tax so measured corresponds to a certain level of redistribution of income, and changes in progression signify changes in redistributive effect; residual progression is the only one obeying this principle.

[7] Atkinson (1970).

[8] In the class of generalised utilitarian SWF defined as above (namely, additively separable, symmetric and concave).

Other measures of the degree of progressivity, linked to its redistributive effect, are the Kakwani index (K), measuring the disproportionality of taxes[9] -defined as the difference between the concentration and Gini coefficients for, respectively, tax liabilities[10] and pre-tax income- and the Reynolds-Smolensky index (RS) of the redistributive effect, given by the difference between the concentration coefficient for post-tax income and the Gini coefficient for pre-tax income. These are overall measures of progressivity and the relation between them is:

$$RS = g / (1-g)\ K,$$

with g standing for the average tax rate that corresponds to the progressive income tax being applied.[11]

Assuming the inexistence of horizontal inequity (and of reranking of tax units), the redistributive effect of a progressive income tax depends not only on the average tax rate, but also on the magnitude of its disproportionality effect.

[9] See Kakwani (1976).

[10] We are referring to concentration curves since the hypothesis of inexistence of horizontal inequity and of reranking, among tax units, induced by the income tax can not be accepted as realistic; therefore, the concentration curves for tax liabilities and post-tax income will not coincide with the corresponding Lorenz curves.

[11] This result is proved by Kakwani (1976) and will be useful in section 4 of this work, when we analyse measures of horizontal inequity.

2) HORIZONTAL EQUITY

The principles of horizontal and vertical equity have a central role in any discussion about income taxation. We have already defined very broadly horizontal equity (henceforth HE) as demanding that people in equal circumstances before tax should be treated equally by the tax system, while vertical equity (VE) requires that tax payments of income units should be differentiated, as a (positive) function of their ability-to-pay, for example in order to equalize their sacrifice on a utility basis. It is essential to develop measures of the extent to which a specific tax system conforms, or not, with these principles, in order to compare systems, analyse the evolution of tax policy in a certain system, or recommend policy changes. We will focus on this aspect in the next chapters of this work; for now, the analysis of the concept of HE and of the justifications for its violation in modern tax systems, will be our concern. Unless we interpret "circumstances" merely as income levels, it already becomes obvious that the principle of HE is related to a tax schedule that cannot naively be formulated as t(x), like we did before.

HE requires that two income units having the same pre-tax equivalent income, the same abilities and the same needs, should have the same tax liability (again, in terms of equivalent income)[12].

Feldstein (1976, p.83) has considered the problem of the diversity of tastes among income units, that renders the typical definition of HE as ambiguous, and has redefined it as follows: "If two individuals would be equally well off (have the same utility level) in the absence of taxation, they should also be equally well off if there is a tax".

Very rarely is the tax unit the single individual, a fact which raises problems in the definition of the utility level of the tax unit, as with equivalence scales. Moreover, utility levels "in the absence of taxation" simply do not exist, unless we assume that taxation does not affect individual decisons (and the shape and arguments of their utility functions), namely in what

[12] It similarly demands, of course, that they should receive the same amount of public benefits.

incentives are concerned. Despite these drawbacks, this definition has the advantage of referring to utility (and not simply to income), requiring that taxes should not alter the ranking of individuals by utility level.

Several views on HE and VE have been expressed by distinct authors, and usually there is less controversy around the concept of HE than around the precise definition of VE. HE is consensually accepted as a fair principle under any modern formulation of the theory of distributive justice, while VE is embodied with disagreements about the "appropriate" pattern of differentiation between different individuals, varying according to normative values. In spite of this, for a long time the analysis of tax equity was primarily concerned with VE and "the problems of HE have received close attention only since the 1930s" (Pechman, 1989, p.42). Moreover, this later principle was commonly seen as a derived consequence from the main rule of VE by many authors, including Musgrave, who, however has more recently[13] changed his view on this, recognizing the normative independency and essentiality of both principles, as fundamental rules of equity on taxation, whose violation leads to a reduction in aggregate social welfare. For social ethics, HE is nowadays not only strongly established, but also established in a more popular and firm way than the principle of VE.

There are three main views on the normative importance of HE[14], directly linked to the role of the state as a redistributor agent through taxation: a libertarian view of social justice (related to authors such as Robert Nozick), in its most extremist fashion, considers the pre-tax distribution of income as the fairest one, HE (as any "distortion" introduced by the state´s activity) being thus a very significant issue. Opposing to this view, the utilitarians (and, also, the advocates of Rawlsian positions) judge the state´s intervention as morally essential in terms of redistribution: HE is not given a great importance, the central concern being with the post-tax distribution of income. Finally, a distinct position focuses not so much on the pre or post-tax distributions of income *per se*, but more on the means by which differences between them are acomplished, HE being morally justified as "a safeguard against capricious discrimination"

[13] For further details, see Musgrave (1990), p.113-114.
[14] Atkinson (1980).

by the state[15] and ensuring "that the law does not serve anybody's self-interest"[16].

The ethical status of HE varies thus according to the view defended by each individual. However, the study of the magnitude of horizontal inequity (henceforth HI) induced by the income tax has always a scientific interest on its own.

For the income tax to respect entirely the principle of HE, it should apply to every source of income and allow for differentiation probably only in terms of tax units' incomes, needs and some special features such as disability or age. The differentiation of tax payments in terms, for instance, of the family composition, normally through a system of allowances and exemptions, is thus a way of implementing this equity principle in real world taxation. However, other types of differentiation are typically used, namely according to the source of income (distinct fiscal treatment of earned income, dividends or capital gains, omission of the imputed value of the rent for home-owners, or of received fringe benefits, exemptions for benefit payments, among other examples), and to the use of income (deductions and allowances are justified for a variety of reasons, such as charitable transfers, medical, childcare and education expenses, insurance premium payments, retirement plans, interest on mortgages). In these two cases we can question the justice argument: they can be seen as ways of arbitrarily discriminating between the income units, leading to violations of the HE principle. Other sources of HI are regional differentiation of otherwise equal fiscal units based on their place of residence (for example: rural *versus* urban areas), as well as problems of tax evasion or of noncompliance of tax payments.

It becomes evident that complexification of the tax code, in terms of the permitted deductions, allowances, credits or exclusions (not directly linked to the household type or composition), will tend to increase the degree of HI among tax units of equal pre-tax income (or comprehensive income) and its simplification will have the opposite effect.

Returning to a utility-based definition of HE, whenever individuals have different utility

[15] Musgrave (1959), quoted on Atkinson (1980).

[16] Lambert and Yitzhaki (1995), p.674. This paper focuses mainly on the legal foundation of the principles of VE and HE as basic features of social justice.

functions, needs and abilities, any income tax will induce HI (assessed in terms of pre and post-tax incomes of tax units); the definition of the relevant needs and family structures that have to be considered in the appropriate setting of the comprehensive tax base itself, embodies value judgements and arbitrary choices. Illustrating this, Manser (1979, p.224) says: "without knowledge of how the household objective function is related to the utility functions of the individuals forming the household, it is not clear how welfare comparisons can be made between married and single individuals", and problems are worsened when we consider other (more complex and varied) types of households (or tax units).

Therefore, we have to stress that a totally horizontal equitable income tax is not possible to design in real world tax systems. The issue is, thus, one of minimizing the degree of HI.

In reality, the heterogeneity of the tax units in terms of their composition, family type and special needs (among other factors), complexifies the analysis of HE, unless of course the tax code requirements would correspond entirely to some universally accepted concept of equivalent income[17]. For this hypothetical situation, HI would arise whenever income units with the same pre-tax equivalent income ended up with distinct post-tax equivalent income. With a social heterogeneous population, this is the typical approach adopted in the analysis and measurement of this issue. However, it still raises major problems in international comparisons for cases where distinct countries consider different equivalent scales in their assessment of HI.

The modelling of the income tax for equity analysis purposes demands then, either the use of equivalent income, a solution which, as we will see[18], if totally meaningful in theorethical terms, has a subjective (and controversial) foundation in its application, or, alternatively, distinct schedules for each truly homogeneous subgroup of the total population (however defined). Only for income units all of the same type analysis, for instance, of the degree of progression is a sensible exercise. Social heterogeinity implies the definition of a set of progressive tax schedules, each being simultaneously the least horizontally inequitable possible, and a global analysis of progressivity and of HI is not in general possible.

[17] The definition of equivalent income and the consideration of the problems related to this concept will be developed in section 3.

[18] Problems with equivalence scales will be treated deeply in section 3.2 a) of this work.

Having said this, the main idea which we want to stress at this point is that the redistributive effect of a progressive income tax (formerly, under the restrictive set of assumptions made, identified with its vertical effect) depends not only on VE considerations, but also crucially on the degree of HI. The unequal treatment of equal tax units by the income tax globally causes a reduction in the magnitude of its redistributive capacity, implying a loss of its vertical performance when compared to a reference situation of a progressive tax which is horizontally equitable. The assessment of the extent of HI is thus essential, and it may be an important factor in understanding why, in fact, income tax systems may be less progressive than *a priori* we would expect.

Developing measures that allow us to confirm the extent to which a tax system is vertically and horizontally equitable, relating these two concepts and understanding their inter dependency, has been a major task in recent decades in the public finance literature. In the next chapter we will focus on some of the practical concepts and methodological problems linked to this issue and a few analytical measures will be surveyed in chapter 4.

3) MEASUREMENT ISSUES AND PROBLEMS

3.1) THE CONCEPT OF HI IN PRACTICE

The issue of defining a methodology to summarize the magnitude of existing HI and its relation to the global performance of the income tax, incorporating simultaneously, in overall indicators, VE characteristics, is complex. This complexity starts with disagreement between distinct theorists on how HI is manifested, some identifying it with reranking effects and others clearly distinguishing it from these: HI in the income tax may change the ordering of individuals from the pre to the post tax distribution of income, although this is not an inevitable effect.

Feldstein, defines HI in terms of the reranking of utility levels between tax units, and he has thus proposed a measure of this effect in terms of the rank correlation coefficient between the ordering of utilities from the pre to the post-tax situation. The largest disadvantage of this view is that it demands the specification of the tax units´ utility function(s), which is highly controversial in practical, empirical terms. Besides this, the concept of HI is intuitively related to, and theoretically founded on, aspects of income redistribution, of income differences that arise in cases where they did not exist before, as a consequence of income taxation; measures based on ranking effects are linked to measures of income redistribution among pre tax equals, but this relation is not simple[19].

[19] The reranking approach to the measurement of HI has emerged since the beginning of the 80s (initially with Atkinson and Plotnick´s papers) as a separate -although partial- strand in this type of literature. We have chosen not to focus deeply on this specific approach; for a detailed comparison of both strands see Lambert and Ramos (1997b).

Atkinson (1980) has shown that inequality in post-tax income is sensitive to ranking changes. He considers that the measurement of HI is related both with alterations in the position of pre and post-tax Lorenz curves, and with the degree of "mobility" caused by income taxation (establishing an analogy of this latter effect with the measurement of social mobility across several generations); reranking effects in themselves do not affect the magnitude of post-tax inequality among pre tax equals, but they can nevertheless constitute an important factor influencing some ways of measuring these inequality effects[20]. His proposed index of reranking effects can be seen as a partial measure of HI. Similarly, a measure exclusively based on redistribution features, and ignoring reranking effects, may not be seen to account for the total degree of HI induced by the income tax.

Plotnick (1981) proposed the Preorder Inequity Index (PII) as a measure of HI that, however, considers only the magnitude of reranking effects caused by the income tax policy. He chooses to ignore aspects of HI linked to inequality (introduced by the tax among pre-tax equals) and focuses exclusively on reranking effects: "while determining the appropriate preredistribution ranking in the absence of a particular redistributive policy presents serious problems, it is ranking which is essential for assessing HI"[21]. PII is constructed on the basis of the normalized relation between the post-tax Lorenz curve and the post-tax concentration curve (with the value zero meaning perfect HE, and 1 implying maximum HI), being sensitive simultaneously to reranking effects induced by the income tax, and to the levels of income differences observed.

Besides being a partial measure of HI, PII has the additional disadvantage of embodying a value judgement that may not however be particularly appealing to social analysts: reranking effects affecting tax units at higher pre-tax income levels are valued more by this index than the ones occuring at lower income levels.

In the beginning of the 1980s, after Feldstein´s and Atkinson and Plotnick´s important papers, the methodological tendency in HI measurement was thus to focus on the assessment of reranking of tax units affected by the income tax. This happened also due to some empirical

[20] For further details on how reranking effects affect the Lorenz curve and the Gini coefficient, see Atkinson (1980), p. 9 to 13.
[21] Plotnick (1981), p.284.

difficulties in measuring "classical" HI due to the problem of the identification of equals (as we will see in section 5). The no reranking condition is, as we mentioned, necessary for the existence of HE but it is not sufficient. Therefore, this measurement approach is fraught with problems and has an additional drawback of implicitly assuming that the pre-tax ranking of tax units by levels of well-being (that should not be "disturbed" by the income tax) is the fair one.

Musgrave (1990, p.117) proposes an abstract notion of "social welfare cost" as a measure of HI: "applied to any one group of equals, HE performance is measured by the excess of the combined actual welfare cost for that group over what it would have been with equal division of liability within the group", and a different local measured level of HI ("excess cost") can be determined for each group. The entire level of HI (somehow also related to the principle of VE) has, then, to be built by some aggregation procedure (which Musgrave considers that should be independent from local levels of well-being[22]) and a value of zero for this index would reflect perfect HE.

This approach, despite having the disadvantage of the concept of "well-being" of tax units not being operational in practical work, is theoretically appealing and does not focus so much on the study of reranking (or tax induced mobility) aspects.

The need to choose the least horizontally inequitable income tax system (or income tax reform) makes indispensible the measurement of this aspect. A currently followed approach in measurement theory, in the line of Musgrave´s proposed measure, is related to the assessment of local HI among pre-tax equal income units, using for this purpose inequality measures, followed by some aggregation procedure of these local measures into an overall measure of HI. This final measure, as we will see, incorporates, either implicitly or explicitly, the "classical" HI considerations and the reranking problems of tax units.

In this context, measurement of HI has all the well-known methodological problems of inequality measurement, that will be the concern of the next section of this work.

[22] We will mention later some problems with aggregation procedures of local HI measures.

3.2) METHODOLOGICAL PROBLEMS IN MEASURING HI AS INEQUALITY

The initial obvious difficulty is the practical delimitation of the groups of equals; using simply pre-tax income as the operational variable for this purpose -"equals" being those income units with the same pre-tax level of real income- is unsatisfactory, as we will see. Defining the tax unit is also problematic and choosing an appropriate measure for inequality (local HI) assessment has distinct implications in normative terms, as the choice of the aggregation procedure (for the determination of an overall effect) has.

Another major problem is related to the definition of "inequity" and the quantitative assessment of its extent. Johnson and Meyer [23] point that we can: "define the magnitude of an inequity in terms of the dollar value of the tax discrimination involved. It can be assumed to be proportional to the value involved or it can be treated as an increasing or decreasing function thereof.....inequities are not all equal but do in some way have different magnitudes". The economic measurement of HI always embodies value judgements, whose implications it is essential to know *a priori*, since they also determine the resulting level of measured inequity, besides the inequity *per se*. We are focusing exclusively on measuring locally HI as economic inequality[24], and by inequality it is meant the existence of differences in the distribution of a relevant variable among a specific population (in this case, the pre-tax equals), in a period of time. The term embodies simultaneously a moral aspect related to the undesirability of its occurrence. The positive vs. normative approaches to measurement of inequality seem clearly diverse, but in practice they are intimately linked as it will be shown.

We will expose the measurement problems as follows: a) prior problems to consider when

[23] Johnson and Meyer (1962), p.458.
[24] Although this form of inequality is usually linked to political or social inequality.

measuring local HI as economic inequality; b) the choice of the inequality measure; and finally, c) the aggregation procedure into an overall HI measure.

a) Measuring economic inequality: "prior" problems

The initial problem is: **inequality of what?** How can we best represent tax units well-being using a unidimensional scale, a unique variable, for this purpose?

The most usual dimension considered when assessing inequality is income (pre and post-tax income to assess local HI)[25] and several factors have to be taken into account when analysing it. In general, its differences must be judged having in mind not only that individuals have distinct levels of earnings and wealth (the two main sources of income), but also that they have different needs, distinct tastes (regarding work/leisure, investment in human capital saving and risk, for instance) and abilities, leading to different choices/opportunity sets, that they face expectable variations in income according to their life-cycle and that they are affected diversily by random factors such as luck. Therefore, equality in pre-tax income should not always entail the inexistence of inequality or injustice *a priori*; the income tax can, thus, be seen as a source of inequality enhancement between "equals" but can also reduce inequity between income units with the same pre-tax income. Taking the pre-tax distribution of income -egalitarian among equals so defined- as the reference in terms of fairness may be misleading. Nevertheless, this is the typical procedure taken as valid in empirical work.

Other dimensions that could be considered are wealth (a stock, at some moment of time, of multidimensonal values -like physical goods, financial assets or human capital- thus hard to assess fully) or some welfare indices, linked to utility concepts -like in Feldstein and Musgrave´s formulations of the concept of horizontal equity- and even harder to measure.

[25] We refer only to current income (abstracting from lifetime income and mobility-type effects in the unequality of equals); equivalent income will be considered deeply below.

Related to this last dimension is the concept of full-income[26], including money and non-money income, where the former includes wage and non-wage (e.g. capital gains) income and the latter refers to the satisfaction (utility) derived from work and leisure, to the value of own production and the services of physical wealth, or, more broadly, to an individual´s (or income unit´s) opportunity set. HE would require that this potential power to consume goods -whether or not it is fully exercised- was preserved by the income tax for each and all groups of units with equal opportunity sets before the tax.

Due to obvious difficulties in operating this concept in practice , empirical studies normally concentrate on income or on equivalent income. According to Simons[27], "income is the value of rights that a person might have exercised in consumption without altering the value of his wealth" or, equivalently, is "the sum of (1) the market value of rights exercised in consumption and (2) the change in the value of the store of property rights between the beginning and end of the period". In this definition are included many variables that, however, are often disregarded in usual tax codes and oficial statistics, like capital gains and losses, the imputed rent in the case of house-owners, the value of production for own consumption and fringe benefits.

In spite of all the loss of accuracy that simplifications imply and considering that with this unidimensional variable, a large amount of relevant information about well-being and tax units´ utility is lost, real money income (pre and post tax) is a very commonly used variable for the assessment of HI, measured as inequality.

The next methodological problem to consider is **inequality of whom?** The problem of the definition of the tax unit is also controversial, since different possible samples, leading to distinct conclusions about horizontal inequity, can be considered: the individual, the nuclear family, the household, or the "equivalent adult". The appropriate unit that the tax code and empirical studies should consider is "a unit in which economic decisions about paid and unpaid

[26] Barr (1993), p.133.
[27] Quoted in Atkinson (1983), p.37.

work and about consumption are made jointly and in which resources are pooled".[28]

Therefore, if we take the individual as the basis of analysis we can reach very biased results, seeing that many individuals with zero (pre-tax) income (e.g. housewives, students) share in fact, income (and consumption) with other individuals. On the other hand, taking the household as a unit implies quite strong assumptions about the way resources are shared in its interior. Generally, the larger the units considered (for any distribution of income across these units), the smaller will be the (global) inequality observed.

[28] Piachaud, in Barr and Whynes (1993), p.108.

The measurement of the degree of income sharing within a unit is then essential, but extremely difficult to achieve[29]. Therefore, general scales of equivalence of income have been constructed, although not without some controversy, for the purpose of comparing units with different numbers (and types) of elements. An equivalence scale is a device to convert incomes of distinct income units into a common base, measuring the purchasing power -or utility- per unit, against which horizontal inequity is then properly assessed. Income *per capita* is considered a misleading indicator , since there are economies of scale within a unit and individuals´needs vary, according for instance to their age. Equivalising income makes then households with different needs and composition comparable for normative purposes, but the choice of an "appropriate" equivalence scale makes a considerable difference to the assessment of the degree of inequality (local HI) of their "well-being".

The starting point may be a single person (with a value of one) and for each additional member different (subjective) ponderations are considered, trying to equivalise income between households of different types; a set of deflators (considering, namely, specific needs and age) is built, and the adjusted income results from the division of the household income by the correspondent value of the equivalence scale.

The theory of equivalence scales arises from demand theory and it typically assumes that all households have equal utility functions (neglecting the special case of the sick, for instance). Moreover, it is inevitably taken as valid the possibility of making interpersonal comparisons of utility among (and within) different units[30].

[29] For further details on the measurement of intra-household inequality, see Borooah and McKee (1994).
[30] For further details on this criticism, see Fisher (1987).

Another central contentious aspect is related to the impossibility of transferring units of equivalent income between different tax units[31]. In real world tax systems, income -and not equivalent income- is the variable used for redistributive purposes, raising severe practical problems when we use the mentioned concept. Linked to this point is also the inconsistency of some of the underlying SWF to the equivalence scales, fact which has led Ebert to refer to the concept of "equivalent adult", when analysing the application of the principle of transfers to equivalent income.[32]

Besides their subjectivity and possible theoretical inconsistency, another interesting criticism made of these scales is that they assume the fact that parents´well-being depends only on their expenditure level and ignores the fact that they may derive utility from their children (i.e. a couple with a child could in fact have a higher level of "full-income *per capita*" than a couple without children).

Tax equity analysis typically abstracts from all these problems assuming a socially homogeneous population. The real world income tax normally disregards (and violates) the utility-based definition of HE. This analytical procedure may however involve major inconsistencies and less accurate measurement results than the use of equivalence. An obvious extension of the utility-based definition of HE to the case of a social heterogeneity requires households with the same "pre-tax equivalent income" to be treated equally, i.e. to end up with the same "post-tax equivalent income".[33]

Finally, we need to ask: **inequality considering which period?** To assess inequality, we can consider diverse possibilities of time period, e.g. 3 months, a year, a decade or an individual´s lifetime. The most commonly used dimension is the year, but the choice depends on the objectives of the specific study. In general, the longer the time period considered, the less

[31]Consider, for instance, the equivalence scale: Z = (Number of adults + ½.Number of children) ^{½} . If we want to transfer 1 unit of equivalent income from a rich single to a poor couple, for the former this 1 unit of equivalent income corresponds to $1, while for the latter 1 unit of equivalent income demands $1.41; so, the problem is unsolvable, and the transfer is not possible!

[32] For further details on SWF underlying equivalence scales and the concept of "equivalent adult", see Ebert (1997).

[33] For further details on comparison of households with different structures for HE analysis purposes, see Manser (1979).

unequal will (seem to) be the distribution of income. For the study of HI, the time period considered is typically the fiscal year, whose definition may, however, differ considerably from one country to another.

All the above-mentioned problems affect the results of the measurement exercise and can raise major problems for international comparisons of tax systems using distinct definitions of income (which is, already by itself, a very limited proxy of units' utility level), income unit, fiscal year or different equivalence scales.

b) The choice of the inequality measure

When we express income inequality by a unidimensional variable, this corresponds to some implicit abbreviation of a SWF, considering it as a multidimensional phenomenon. The nature of the inequality measure is therefore dependent on the properties and characteristics of an

original, global SWF, whose features it is important to consider and analyse[34]. A further previous observation is that any inequality measure uses, as an ideal reference, the totally egalitarian distribution of income; for the study of HE in the income tax, this may be seen as "correct" only when we are using equivalent income and assuming identical tastes and abilities among the various income units.

The issue is thus how can we best construct a scalar measure of inequality, and what normative judgements and views about inequity underlie it?

Inequality indices have typically been developed with major attention being placed on statistical rather than on ethical purposes. In any case, a notion of equity, either implicit or explicit, is always embodied in these measures, and many statistical indices have proved to have undesirable properties from a normative perspective. However, a convincing welfare rationale must always be present in any statistical index to be used in inequality analysis. Contrasting with this approach, explicitly normative inequality measures have been built *ab initio* from an ethical point of view, departing from the formulation of some convincing and consensual SWF, such as the Atkinson index, as we will see below.

In this context, we will present some possible indices for the measurement -as inequality of post-tax income among pre-tax equals- of HI, dividing this presentation between: i) statistical measures (or implicitly ethical measures), and considering, for these, both the concepts of relative and absolute inequality, and ii) normative measures (or explicitly ethical measures).

i) Statistical measures

We initially need to distinguish between the concepts of relative and absolute inequality. The former means that inequality (both the feeling of fairness and the way to measure it) is

[34] On this aspect see, for instance, Kondor (1975).

invariant to a proportionate change in all income levels (i.e. when income shares remain the same), while, alternatively, the latter implies that inequality is invariant to equal additions of income to all income levels.

Statistical measures of **relative** inequality should satisfy a set of desirable properties, besides having a convincing welfare rationale. These properties are:

- scale independence, demanding invariance of the inequality index to equiproportionate changes in all levels of income;

- symmetry, meaning that permutation of incomes in the distribution does not alter the value of the index;

- obeyance to the principle of transfers, implying that the index must be directly sensitive to (rich-to-poor) transfers in the distribution (that do not change previous rankings);

- decomposability, demanding that overall inequality should be decomposed, by the use of the index, in terms of inequality within and inequality between subgroups of the total population of income units. If we wish the index to display this property, this will restrict the available choice of indicators; for instance, the mean logarithmic deviation conforms to it, but the Gini coefficient fails to obey it completely when there are overlaps between subgroup income ranges[35].

If an index of inequality is decomposable, overall inequality (I) can be expressed as a weighted sum of the inequality values for the inequality within each subgroup of the population (with the weights depending on population shares and/or income shares), and another term measuring inequality between groups (I_B) assuming that in each group every unit has the correspondent mean income for that group: $I = I_B + \Sigma_k \alpha_k I_k$, where I_k stands for inequality within each of the subgroups in the population and α_k are the corresponding weights for overall inequality within subgroups.

[35] In fact, this is one of the motives why the use of the Gini is so popular in HI measurement (including reranking effects), as we will see.

The mean logarithmic deviation (MLD) is a statistical measure which has been used in HI measurement[36], mainly due to its decomposability properties as an inequality index. Mathematically, for a certain income (or equivalent income) distribution $\{X_1, X_2,....X_n\}$, the MLD is given by: $M = \Sigma \ln (\mu / X_i) / N$; in the case of the measurement of HI, each group of incomes X_i will correspond to the post-tax incomes of those who were equals before the tax (i.e., in this case, had the same income -or equivalent income- before the income tax). The MLD belongs to the general entropy class of inequality indices, thus being (completely decomposable) which means that aggregate inequality (overall HI) can be expressed as a function of the subgroup inequalities (related to local HI), of mean incomes in each subgroup (to determine inequality between subgroups) and population shares[37] (see above). This family of indices -the entropy indices- is usually strongly favored on the basis of its statistical advantages, although the study of its normative content is still quite incipient and its welfare rationale remains unclear, reason why they are not consensually favored on ethical grounds.

A widely used measure of relative inequality is the Gini coefficient, which summarises into one number the information embodied in a Lorenz curve (since it corresponds to twice the ratio of the area between the Lorenz curve and the diagonal and the total area under the diagonal, or line of complete equality). Mathematically, the Gini is given by the formula:

$$G = 2.COV (X , \text{Rank of } X \text{ in the income distribution}) / \mu ,$$

with X standing for the (equivalent) income levels (preordered in ascending order) and μ for the mean income. Alternatively, the Gini is half the arithmetic average of the absolute differences between all pairs of incomes:

$$G = \Sigma_i \Sigma_j \mid Xi - Xj \mid / 2.N^2.\mu ,$$

where N stands for the total number of income units. G will vary between 0 (for total equality)

[36] See, for instance Lambert and Ramos (1997) for an application to the Spanish income tax system.

[37] For this index, the aggregation weights for the inequalities within groups are a function only of population shares and thus are not dependent on income shares, differently from what happened with the Gini, as we saw before.

and 1 (for maximum inequality). The Gini uses all the information available for the distribution and normalises it, being scale independent, emphasizing absolute differences in units´ incomes more than relative differences, and obeying the principle of transfers. However, the implicit SWF in the Gini coefficient may be considered as not specially appealing since the weights of people's incomes in the distribution are determined by the person´s (or other type of income unit) rank-order in the distribution and it does not satisfy the principle of diminishing transfers. The SWF implied in the Gini has no convincing ethical foundation based on desirable properties such as additivity and strict concavity departing from individualistic preferences. The Gini is not an utilitarian-based inequality index.

Nevertheless, Sen (1973) has shown that a utilitarian rationale could be provided for the Gini coefficient, with his "pairwise maximin criterion": "suppose the welfare level of any pair of individuals is equated to the welfare level of the worst-off person of the two. Then if the total level of the group is identified with the sum of the welfare levels of all pairs, we get the welfare function underlying the Gini coefficient".

Moreover, Lambert (1985) has also established a convincing welfare rationale where the Gini fits, based on the concepts of deprivation and altruism.

Concluding, there are "valid" normative rationales for the use of the Gini coefficient as a measure of relative inequality, besides the statistical advantages pointed out above. The fact that the Gini does not satisfy completely the decomposability property when the income ranges overlap, works, paradoxically as an advantage when measuring HI; in this situation, the overall Gini for post-tax income (corresponding to the post-tax concentration curve) can be expressed as: $G = G_B + \Sigma_k \alpha_k G_k + R$, with G_B accounting for inequality between (pre-tax equals) subgroups (assuming that in each subgroup there is HE), G_k measuring inequality within subgroups (i.e. "classical" HI), with α_k being weights function of population and income shares, and, finally, R standing for a general measure of the extent of overlap between subgroup distributions[38], in other words, measuring the reranking effect[39].

[38] In the case of a post-tax income distribution when the subgroups are pre-tax equals groups, if we decompose the Gini, R amounts to an Atkinson-Plotnick index of reranking.

A second type of statistical measures are those related to **absolute** inequality. If the social analyst believes in absolute inequality, an index to assess it demands, instead of scale invariance, the property of invariance to equal additions to all incomes: equal additions of income to all income levels do not change the value of absolute inequality[40]. The most commonly used index of absolute inequality is the absolute Gini, defined as: $AG = \mu.G$, (where G stands for the normal Gini - index of relative inequality), varying between 0, for complete absolute equality, and μ, for maximum absolute inequality.

The idea of absolute inequality, in measurement theory and in ethical terms, is usually far less appealing than the concept of relative inequality, reason why we are not concentrating so deeply on it. In any case, this concept can also be applied to the study of HI, as we will see in further detail in section 4.

ii) Normative measures

[39] For further details on the mathematical decomposition of the Gini index, see Aronson and Lambert (1994).
[40] Note that this would reduce relative inequality.

Statistical measures implicitly embody arbitrary SWFs, implying distinct value judgements (sometimes unacceptable, violating essential normative principles) concerning inequality. Therefore, initially Dalton, and then Atkinson (1970), proposed measures that reveal these values explicitly; only proceeding this way can it be clear what objectives are being incorporated in the analysis as a result of adopting a certain measure. We will be dealing exclusively with the concept of relative inequality and, again, only income information is considered ethically relevant for the assessment of inequality (and, consequently, of HI)[41]. In this section we will focus on the Atkinson index[42], which can be used in the assessment of the magnitude of HI, as it will be further illustrated in section 4 of this work.

The SWF explicitly considered in the formulation of Atkinson´s index is utilitarian-based, an average of individual utility-of-income functions, being additively separable, and symmetric[43]. Moreover, the class of utility functions allowed is restricted to the ones that are increasing (Paretian) and concave (i.e. transfer-approving). Atkinson put forward the idea of "the equally distributed equivalent level of income" (EDE, hereafter): "the level of income *per* head which if equally distributed would give the same level of welfare as the present distribution"[44], leading to the index of relative inequality, defined as:

$$I = 1 - EDE/\mu ,$$

varying between 0, for complete equality, and 1, for maximum inequality. Relativeness of inequality is assured by a constant degree of (relative) inequality aversion, ε (ε>0 to ensure

[41] We can use real income or, alternatively, equivalent income if we desire to incorporate different needs in our analysis, making it more realistic and consistent with real world social heterogeinity.

[42] A possible topic for future research could be linked to the study of the decomposability properties of the extended Gini index, developed by Yitzhaki (1983) with a very interesting set of normative properties and some similarities with the Atkinson index, and its eventual application to the study of HI. This task will not be performed in this work.

[43] We can consider these general properties without the need of specifying a particular functional form for the SWF.

[44] Atkinson (1970), p.250.

transfer approval), corresponding to a general utility function having the form:

$$U(X) = A + B\,[X^{1-\varepsilon} / (1 - \varepsilon)]\ ,\ \text{if}\ \varepsilon \neq 1 \quad \text{and}\ U(X) = \log_e (X),\ \text{if}\ \varepsilon = 1$$ [45].

(A and B>0 are constants)

It becomes now clearer that a value of 0 for I can be observed in two cases: either we have an equal distribution of income (i.e. $X_i = \mu$), or $\varepsilon = 0$ (situation of indifference to the distribution of income). As X_i diverges from the average value or the degree of inequality aversion (chosen by the social analyst) increases (with comparatively more weight being given to transfers at the bottom than at the top of the distribution), the value of I will rise. When $\varepsilon = \infty$, society´s concern is assumed to be exclusively with the worst-off individual (corresponding to a Rawlsian view of social justice).

The idea of "inefficiency of inequality" is central in Atkinson´s formulation: inequality (and so, as we will see, HI) wastes social welfare. For a given ε, the *per capita* cost of inequality can be represented by: $C = (\mu - EDE)$; if a value of C *per capita* were "thrown away" and the remaining income were redistributed equally, the level of welfare would be the same as with the actual, unequal, distribution of income. The value of C obviously varies directly with the degree of inequality aversion.

The Atkinson measure is thus theoretically very appealing but, "just as the earlier measures were not normatively much use, so this measure is not descriptively much use"[46]. International comparisons of data on inequality may be meaningless, unless we hold ε constant for all countries, through time.

[45] Atkinson´s index is formulated from an analogy between risk theory under uncertainty (namely, risk aversion measures) and inequality theory.
[46] Culyer (1991), p.123.

Moreover, Sen (1982, p.419) argues that this approach ignores the distinction between inequality and losses in social welfare from inequality; consequently, I should be seen as a measure of an analyst's sensitivity to inequality and not of inequality itself[47]. Atkinson considers that this does not diminish the validity of its use, since, it "honestly" (by contrast with statistical measures) reveals the difference between description and social judgements on welfare. In fact, for a certain distribution of income, and given ε, $W = \mu(1 - I)$ is a welfare indicator, known as the money-metric welfare indicator.

The indices of relative inequality that feature in recent HI measurement analysis are the Gini coefficient, the generalized family of entropy measures (of which the MLD is an example) and the Atkinson index. For the assessment of absolute inequality, the use of the absolute Gini coefficient is one possible choice.

[47] For further details on normative problems, see Sen (1982).

c) The aggregation procedure

A final methodological stage in the measurement of HI as inequality, is related to the aggregation method. We have seen how, after previous indispensible definitions (of the relevant variable, income unit and time period), distinct inequality measures, with specific statistical properties and normative features, can be used to assess local HI among (previously defined) pre-tax equals, considered in homogeneous groups. The last stage of this exercise is thus to aggregate these local measures of HI for the whole population of tax units, into an overall measure of the HI induced by the income tax system. This global measurement involves again value judgements. Besides the ones related to the chosen measure by itself, there is some polemic around the issue of using a pure (i.e. non-dependent on income) or an impure (i.e. income-dependent) weighting scheme for this purpose; the social analyst´s ethical views on the relative importance of local horizontal inequities at distinct levels of post-tax income will underlie the definition of the precise weighting scheme. Therefore, aggregation of horizontal inequities can be made either independently of income levels, depending only on population shares (the relative sizes of the the groups of pre-tax equals), or weighting differently these inequities for different levels of income; for instance, greater importance could be attached to local horizontal inequities among the groups with the lowest incomes than among the groups of pre-tax equals with the highest levels of income[48]. By contrast, in the former case, of pure aggregation, we are implicitly assuming that: "any inequity of a given size is worth equal weight independent of the proportion of income or tax liability that the inequity represents"[49].

Musgrave (1990, p.117/8) defends the use of a pure weighting scheme as a way of avoiding "inappropriate comparisons between unequals", but this issue is not consensual, as it is clear by what was said above.

In any case, as with the choice of the inequality measure for the local assessment of HI, the analyst must select a preferred procedure having present the normative implications of doing so.

[48] We have seen that for the MLD the weighting scheme is pure, depending only on population shares, while for the Gini coefficient the weights are dependent on population shares and income.
[49] Brennan (1971), p.454.

Summing up, having determined which particular proxy of utility is to be evaluated (which includes, usually, selecting a particular equivalence scale to apply to pre and post-tax income), the assessment period, the type of tax unit to be considered, the inequality measure to be used in the measurement of local HI, and the aggregation procedure to use in the evaluation of an overall measure of HI, the analyst is then ready to (compute and) examine the relevant income distributions, in our case, the pre and the post-tax income distributions, and to draw some conclusions about the degree of HI induced by the income tax for that population, in that specific period. Evidently, distinct combinations of selected measure/variable/equivalence scale/period/tax unit will lead to distinct conclusions about this degree and about its normative significance, and this is why methodological issues are worth so much concern.

4) ANALYTICAL DESCRIPTION OF MEASURES OF HI AND IMPLICATIONS FOR THE OVERALL PERFORMANCE OF THE INCOME TAX SYSTEM

In this section, we will present different measures of HI, always measured as inequality, and considering for the case of the Gini coefficient also reranking effects, with an Atkinson-Plotnick-type index[50]. For practical purposes, pre-tax equals are assumed to be the income units with the same pre-tax equivalent income (represented by X) and HE would prevail only if these units payed exactly the same tax liability, again in terms of equivalent income. Using real income deflated by an equivalence scale as a proxy of "well-being" makes distinct units (in a social heterogeneous population) comparable for normative purposes, as we saw in section 3[51]. Each group of equals so defined is denoted by S(X), at the level X of equivalent income. The tax unit can be the household, the nuclear family, the family or the "equivalent adult", depending on the specific study and tax code that an analyst considers, and the assessment period is typically the fiscal year.

The different measures that we will consider -the Gini, the MLD, the absolute Gini and Atkinson's index- embody different value judgements on the importance of inequities, on the concept of inequality (relative vs. absolute) and imply distinct aggregation procedures. The link with the vertical performance of the tax when we have HI will now be made clearer, for the study of its implications in the overall performance of the tax system. It is necessary to define

[50] The measurement of reranking effects has an important and independent place in the analysis of HI, having emerged in the 80s with Atkinson and Plotnick´s works and being furtherly developed with the definition of families of mathematical indices, namely by King, Cowell and Jenkins. This specific strand is not covered in our work but is deeply treated in Lambert and Ramos (1997b), who also present simulation results exposing (ordinal) similarities and (cardinal) differences in the way indices of the different types rank tax changes effects.

[51] The use of real income could only be justified if we assumed that all tax units had equal tastes, needs and abilities. The assumptions underlying the use of equivalent income, although debatable and polemic, are not as strong and unrealistic as these and allow a more accurate assessment of HI.

the schedule t(X) (for X>O), for the progressive income tax, as the value of the average tax payment of the units in S(X); if it were applied indistinctively to all the tax units in each group of pre-tax equals, this tax would be totally horizontally equitable and raise exactly the same fiscal revenue as the actual, horizontally inequitable, income tax.

i) Statistical measurement of HI

We will start by considering two statistical measures of **relative** inequality to assess local and global HI, the Gini and the MLD; they imply that any proportional deviation from the mean in tax units´post-tax incomes in each S(X) is valued as the same amount of HI for all pre-tax levels of income.

The **Gini** index has been widely used in the study of HI. Initially, this was done in such a way that it was not possible to distinguish "classical" horizontal inequity effects from the distinct reranking effects[52]. It is clear that the income tax may lead to changes in the ordering of the tax units from the pre to the post-tax distribution of income: units who were in the same group of pre-tax equals may be in distinct post-tax ranges; therefore, measures that are built based on pre-tax rankings will typically understate the true extent of inequality in the post-tax income distribution[53].

We have already mentioned the well-known failure of the Gini index to obey to the decomposability property of global inequality into within and between subgroups inequality, when there is overlapping of the post-tax income ranges. Following Aronson, Johnson and Lambert (1994) a realistic schedule of tax payments, for any income unit with a level X of equivalent income, can be established as: $t = t(X) + \theta$, with t(X) as before and θ being a disturbance with zero mean for each level of equivalent income, that reflects precisely the

[52] See, for instance, Kakwani (1984) where the total redistributive effect of taxation is decomposed in terms of HE and VE but, in fact, ignoring the classical unequal-treatment-of-equals effect, measuring only the effect of reranking.

[53] Atkinson (1980) showed that the Gini coefficient of post-tax income is reduced as a consequence of reranking effects of tax units considering post-tax income.

existence of "classical" HI for each level X. The income tax, however, can also cause the reranking of these tax units as a consequence of this departure, and this is essentially a distinct effect since it occurs between unequals.

Using the Gini, we had the general expression: $G = G_B + \Sigma_k \alpha_k G_k + R$, and we can thus express the Gini for the post-tax income distribution, decomposed across the groups S(X) of pre-tax equals, as:

$$G_{X-T} = G_B + \Sigma_x \alpha_x G_{S(X)} + R ,$$

where G_B, the Gini for between subgroups inequality, measures the inequality that would exist if every tax unit had payed income tax totally according to t(X), i.e. assuming that in each subgroup of equals S(X) all units bear an average tax payment, *as if* no HI existed; G_B reflects thus a vertical feature of the tax system. The indices $G_{S(X)}$ measure inequality within each group S(X), among pre-tax equals that became unequals as a consequence of income taxation (local HI), with α_x representing the weights used for their aggregation; these weights result from the product of population and post-tax income shares of income units in each S(X), and therefore are "impure" (income dependent) weights; the total value $\Sigma_x \alpha_x G_{S(X)}$ is thus a measure of "classical" HI. Finally, R, the residual, corresponds in fact to an Atkinson-Plotnick type index of reranking, measuring the magnitude of this effect as a result of the existence of inequality of treatment of tax units in each S(X)[54].

To relate HI so measured with the overall performance of the income tax system, we can analyse its effect on the redistributive effect of the tax using the Reynolds-Smolensky index of the redistributive effect (RE), i.e. of the reduction in overall income inequality as a result of progressive income taxation. We have: $RE = G_X - G_{X-T}$, where G_X stands for the Gini index for pre-tax income[55]. Considering G_{X-T} defined as above, the redistributive effect can therefore be written as: $RE = (G_X - G_B) - \Sigma_x \alpha_x G_{S(X)} - R$;

[54] The Atkinson-Plotnick index of reranking is given by: $(G_{X-T} - C_{X-T})/2G_{X-T}$, with C_{X-T} being the concentration coefficient for the post-tax income distribution ordered by pre-tax ranks, while $R = G_{X-T} - C_{X-T}$ therefore very similar to the former.

[55] As we saw in section 1, when the income tax is progressive, $G_X > G_{X-T}$.

The term in brackets can be interpreted as the vertical effect of the tax (V), its contribution to inequality reduction that would be verified if no unequal treatment (in terms of equivalent income) of tax units existed. As we saw, the term $\Sigma_x \alpha_x G_{S(X)}$ represents "classical" HI (H) and R, measures the extent of the reranking effects. Rewriting the expression, we have:

$$RE = V - H - R ,$$

with the redistributive effect of the tax (its overall performance) being negatively affected both by the existence of (classical) HI and of reranking of tax units.

In terms of welfare impacts, this reduction in the overall performance of the tax (comparatively to a reference situation with perfect HE) can be evaluated using Sen´s welfare index[56]: $W = \mu (1 - G)$, with μ being average income and G, the Gini. We thus have W_{X-T} as the welfare index for the distribution of post-tax equivalent income and $W_{X(1-g)}$ can stand for a measure of the welfare level that would prevail after the application, to the same pre-tax distribution, of an equal-yield proportional tax. It is immediate that:

$$W_{X-T} - W_{X(1-g)} = \mu (1 - g).RE = \mu (1 - g).(V - H - R), \text{ (g is the average tax rate)}$$

or, considering the relation between the Reynolds-Smolensky and the Kakwani index of progressivity[57], we also have: $W_{X-T} - W_{X(1-g)} = \mu [g. K - (1-g) (H + R)]$, being obvious that: "the horizontal and reranking terms *both* provide subtractions from the welfare superiority of the actual tax code over a flat (distributionally neutral) one"[58].

The **mean logarithmic deviation** (MLD) enjoys the property of perfect decomposability across the subgroups S(X) of pre-tax equals, allowing us to express global inequality of the distribution of post-tax income as a weighted sum of inequality within groups and inequality between groups (the inequality that would prevail if the tax were perfectly horizontally equitable, i.e. if in each subgroup S(X) all income units payed the income tax according solely with the progressive t(X), instead of with $t = t(X) + \theta$, as explained above). For this measure,

[56] Sen (1973).
[57] Please, see the end of section 1 of this work.

the weighting system is not dependent on income levels, only on population shares in each S(X).

Therefore, having $M_{S(X)}$ as the MLD for the subgroup S(X), measuring local HI, M_H can stand for the index of global HI introduced as a consequence of income taxation: $M_H = \Sigma_x p_x M_{S(X)}$, with p_x representing the appropriate weights: population shares for the subgroups S(X) of pre-tax equals.

To assess the implications of HI in the overall performance of the income tax system, in terms of its impact in global inequality reduction, we need to consider also the "inequality between groups" effect: if in each subgroup S(X) every income unit payed the average tax value for that subgroup, there would prevail perfect HE and a measure of the vertical effect of the tax could be obtained under this reasoning, also using the MLD: we can call it M_B (for inequality between subgroups, in the hypothetical situation when no HI is present).

The global inequality prevailing in the post-tax distribution of equivalent income is thus: $M_P = M_B + M_H$, and the redistributive effect of the tax, generally defined as the difference in inequality from the pre to the post-tax equivalent income distributions, comes: $RE = M_{PRE} - M_P$ (with M_{PRE} standing for inequality in pre tax equivalent income measured by the MLD), or:

$$RE = (M_{PRE} - M_B) - M_H ;$$

[58] Aronson, Johnson and Lambert (1994), p.266.

The existence of global HI represents a loss in the vertical performance -defined as (M_{PRE} - M_B) - of the progressive income tax, and a reduction in its redistributive effect. A higher level of redistribution of income could be achieved by income taxation, for exactly the same tax revenue, if the unequal treatment of equals (measured by M_H) were removed from the system, leading not only to a greater inequality reduction, but also to welfare gains for the tax units in each and all S(X)[59]. HI worsens the overall performance of the income tax.

We will now focus on a statistical measure of **absolute** inequality to assess local and global HI, the absolute Gini; this implies that a deviation of a fixed amount of income from the mean in tax units´post-tax incomes in each S(X) is valued as the same amount of HI for all pre-tax levels of income.

$AG_{S(X)}$ will denote the measure for local HI: inequality in post-tax equivalent income among the members of each S(X), that were defined as equals before the tax. The aggregation procedure into an overall index of HI involves, again in this case, a pure weighting system: $AG_H = \Sigma_x p_x^2 AG_{S(X)}$, with p_x defined as before.

The overall performance of the tax can be assessed considering the gain in welfare of the progressive income tax system over the proportional (equal-yield and distributionally neutral) alternative, given by:

$$W_{X-T} - W_{X(1-g)} = \Delta W = \mu (G_{PRE} - G_{X-T}),$$

with μ as the average post-tax income and G_{PRE} and G_{X-T} standing for the (relative) Ginis for pre and post tax equivalent income distributions. If there were no HI, the change (comparative gain) in welfare would have its origin solely on the vertical redistribution effect, i.e. $\Delta W_{VR} = \mu (G_{PRE} - G_B)$, with G_B denoting the (relative) Gini measuring inequality between subgroups as before; therefore, the global (smaller than this) positive variation in welfare comes:

[59] Atkinson (1970).

$$W = \Delta W_{VR} - AG_H - \mu R ,$$

illustrating again that if HI (measured as absolute inequality) and reranking were eliminated from the tax system a higher welfare level could be attained (for the same tax revenue level as in the actual system) by the income units in all groups S(X)[60].

ii) Normative measurement of HI

This topic is related to the Atkinson index. Here, HI is considered as a social cost: a certain amount *per capita* of post-tax equivalent income could be "thrown away" in each group S(X) in order to have perfect HE, with no variation in the actual level of social welfare[61]. The *per capita* cost of inequality will stand as the local measure of HI, call it $C_{S(X)}$: "the *per capita* income saving among the persons in S(X) that would result from equalizing their post-tax incomes with no loss of social welfare"[62]. There is an implicit comparison between the actual post-tax distribution of (equivalent) income and the distribution that would be observed if all these income values were made equal at a lower value (corresponding to the EDE as we saw before, and denoted for instance by ξ) than the average post-tax income for that group (say, $\mu_{S(X)}$), due to the normative meaning of "the cost of inequality"; this cost - standing as the

[60] We are following the formulation presented in Lambert (1995); however, some inconsistency may be detected in this definition of the welfare gain ΔW, since its rationale lies in relative, not absolute, inequality.

[61] The extent of HI so measured will depend thus not only on relative inequality of post-tax income among pre-tax equals, but also on the selected degree of inequality aversion.

[62] Lambert (1995), p.5/6. Note that we can interpret "*per capita*" as "*per* tax unit", since we may prefer not to restrict the income unit to the limiting case of the individual, particularly when we are dealing with equivalent income, as it was assumed; however, for simplicity of exposition, we can refer to income, instead of equivalent income.

measure of local HI - is thus given by: $C_{S(X)} = \mu_{S(X)} - \xi_X$.

An overall index of HI can again be built, using in this case a pure weighting system:

$$C_H = \Sigma_x \, p_x \, C_{S(X)} \text{ , with } p_x \text{ defined as before.}$$

To assess HI's impact in the overall performance of the income tax system, we can start by defining this latter concept as the variation (reduction) in the cost of inequality that the tax system allows: $\Delta C = C_{PRE} - C_{X-T}$, with C_{PRE} standing for the cost of inequality in the pre-tax income distribution and C_{X-T} denoting the (assumed smaller) cost of inequality after the income tax. A hypothetical tax schedule, that would imply removal of HI welfare-neutrally in any group S(X), has to be defined as the locus of the points (X, ξ_X) for all the income range: $t_{WN}(X)$ is the welfare-neutral horizontally equitable tax schedule[63]. The index C_H, of global HI, measures precisely the departures from this schedule, in terms of the extra tax yield per capita (in the whole population of tax units) that could be obtained if the inequitable system were replaced by $t_{WN}(X)$. The vertical features of the tax are then captured by confronting it with this latter schedule; if the cost of inequality in the post tax income distribution after the hypothetical application of $t_{WN}(X)$ is denoted by $C_{WN(X)}$, the vertical performance (VP) of the tax can be defined as: $VP = C_{PRE} - C_{WN(X)}$[64]. Having so defined HI and VP, the overall performance of the tax system (the achieved reduction in the cost of inequality from the pre to the post-tax distributions of income) can finally be determined by:

$$C = C_{PRE} - C_{X-T} = (C_{PRE} - C_{WN(X)}) - C_H \text{ , or } \Delta C = VP - C_H \text{ ,}$$

illustrating that HI (as in the previous measurement systems considered) causes losses in the overall performance of the tax. A revenue gain would arise from the elimination of HI keeping the welfare level constant.

Concluding, we saw how the removal of HI would improve the performance of the income tax

[63] Following the terminology in Lambert (1995).

[64] Note that, for the previous measures, vertical performance was related to a constant revenue (welfare improving) replacement of the tax schedule by a horizontally equitable tax, while in this approach, VP is assessed by comparing the inequitable tax with one which would be equitable but welfare neutral (revenue improving).

system (in terms of conferring it an extra capacity to reduce inequality - comparatively increasing welfare - or to increase tax revenue - keeping social welfare constant); in the cases of statistical measures of relative and absolute inequality, this is assessed by comparing the actual schedule with a hypothetical revenue neutral (welfare improving) horizontally equitable income tax, while in the case of the (explicitly) normative measurement a welfare neutral (revenue improving) horizontally equitable schedule is considered as the appropriate reference. In each situation, the particular meaning of the measures of HI was, accordingly, established, but a common and crucial conclusion is that its existence *per se* implies a loss in the global performance of the income tax.

5) EMPIRICAL IMPLEMENTATION

When we are dealing with empirical work, having selected a particular database, besides all the methodological problems that were discussed in section 3.2, a further difficulty may be the "identification problem", particularly if we are dealing with relatively small samples: in these there is usually a very small number (if any) of "equals" in each S(X), however defined (in spite of the fact of existing in a much larger number in the overall population); this may lead to an underestimation of HI when measured from the sample so constituted and, in the extreme case of inexistence of exact "equals" in the sample, no HI would be measured with the sample data, although existing in the population of tax units.

This "identification of equals" problem was typically regarded as making impracticable the measurement of "classical" HI, favoring the reranking approach for its assessment, related mainly to the works of Atkinson (1980) and Plotnick (1981) and, as we saw, a very partial methodology.

Nowadays, the identification problem has been substantially overcome, and methodologies including "classical" HI and reranking effects have been established and developed[65]. To deal specifically with this problem, Lambert (1995) proposes to create bands of pre-tax equivalent incomes (using small bandwidths[66]), i.e. groups of "close equals", measuring then - using, for this, the same type of indicators as we referred to in section 4[67] - pseudo-horizontal inequity (PHI) and pseudo-vertical performance (PVP) to obtain a slightly modified version of the magnitude of the redistributive effect of the income tax: "for the interpretation of these pseudo-concepts, it is as if horizontally, the tax acts to increase inequality within close equals groups

[65] Remember the analysis done above for the Gini, or see Aronson, Johnson and Lambert (1994).
[66] The bands can be defined by income ranges -as in Lambert and Ramos (1997)- or by quantiles -the approach that we will follow- which makes for much simplified empirics than used there.
[67] For the mathematical definition of these "pseudo-indices", see Lambert (1995), p.17/18.

and vertically, it acts to reduce inequality between close equals groups"[68].

[68] Lambert (1995), p.16.

Simulation exercises[69] have proved that measures of PHI and PVP (in small samples) tend to the true values of HI and vertical performance (of the whole population of income units) as the bandwidths, in each group of close equals, are narrowed. Moreover, the act of sampling, by itself implying loss of information, was shown not to cause serious loss of accuracy in the measurement exercise.

Therefore, this new methodology allows us to cope with the identification problem when using sample micro data in a serious way, leading to accurate estimates of HI and overall performance of the income tax. We have applied it to an Israeli database[70], with the Family Expenditure Survey data for 1992 of the Central Bureau of Statistics. The original data, considering 5212 households of different compositions as the relevant tax units, included eight variables: weights, number of adults in each household, number of children, pre-tax income, social security benefits, social security taxes, income taxes, and a dummy variable accounting for workers ´ status (salaried vs. self-employed); all income values were in terms of money income, in "shekels", the Israeli currency. In our empirical work the weights were ignored (since the study of the mathematical implications of their consideration for the indicators that were computed is still not sufficiently clear), as well as the dummy variable for ease of computation.

The first step was to determine the money value of "pre-tax income plus benefits", and also of "pre-tax income plus benefits less direct taxes (which include both social security and income taxes)"; then, the households´ money income values of pre-tax income and these two extra variables were transformed into "utility" (equivalent income), using the equivalence scale[71]:

$$Z = (\text{Number of adults} + \varphi.\text{Number of children})^{\theta},$$

with φ accounting for the special needs of children, and θ for the economies of scale within the household (we defined $\varphi = \theta = \frac{1}{2}$).

[69] For further details on this, see Lambert (1995) or Lambert and Ramos (1997).
[70] The author tried to get data from her own country, Portugal, but unfortunately her request to the appropriate department of the Portuguese Ministry of Finance was denied.
[71] Following the procedure adopted in Aronson, Johnson and Lambert (1994).

Proceeding this way, three new variables were constructed: X, standing for "pre-tax equivalent income", XPB as "pre-tax equivalent income plus benefits", and XPBMT referring to "pre-tax equivalent income plus benefits minus direct taxes". After sorting this sample by X, the last biggest 12 observations[72] were deleted to facilitate the calculations, more specifically to allow the division of this data into 100 subsamples (100 percentile groups) of 52 observations each- 100 groups of close equals, in which the observed values were averaged out.

The intention of our study was to measure and analyse the redistributive effects of benefits and direct taxes, as well as the magnitude of "classical" HI[73] and reranking of tax units induced by them, in three possible sets:

a) set 1: when comparing X with XPB, i.e. considering the impact of (equivalised) benefits on the distribution of pre-tax equivalent income;

b) set 2: when comparing XPB with XPBMT, i.e. considering the impact of direct taxes on the distribution of pre-tax equivalent income plus benefits;

c) set 3: when comparing X with XPBMT, i.e. considering the overall impact of the net system on the distribution of pre-tax equivalent income.

In this, the transformatiom from gross to net income is considered to take the form:

$$X \rightarrow X+B \rightarrow X+B-T,$$

whilst plainly $X \rightarrow X-T \rightarrow X-T+B$ is another possibility. The intermediate results will, of course, be path-dependent. The chosen sequence indicates that benefits may be taxable, whilst the latter, that benefits would be assessed on a claimant´s net income. In practice, there are benefits of both kinds in almost any real world tax-benefit system (including the Israeli one).

[72] In empirical work, the highest observations in terms of (equivalent) income are typically considered as non-representative, so this procedure has no serious implications in terms of the analysis of results.

[73] In fact, what we indeed measured was PHI and PVP, but we will refer to the general concepts for simplicity.

Intuitively, we would expect the benefit and the direct tax system both to be progressive (implying positive redistributive effects in all sets). In terms of HI, it seems logical that in set 1 it should tend to be greater at the bottom of the distribution, while in set 2 it should be greater at the top, and in set 3 HI should be spread amongst the whole population; graphically, this particular observation tends to be confirmed by the plotting of the data for these three cases (see the first three graphs in the annex). Moreover, when going from set 2 to set 3, we could expect HI to be reinforced, i.e. the direct tax system acting to increase the magnitude of HI already induced by the benefit system[74].

The empirical analysis was performed using the Gini index, seing that it allows for the decomposition of the redistributive effect (RE) in terms of vertical redistribution, HI and reranking[75]; concretely, we wished to determine: $RE_i = V_i - H_i - R_i$, with i=1, 2, 3, corresponding to each of our three sets.

Using the SPSS computer package, we determined several covariances between the variables, as well as between the previously determined 100 mean values, determining all the relevant values for Gini indices and concentration coefficients that allowed us to determine the above-mentioned decomposition of the RE for the different sets. RE_i was computed from the global sample of 5200 observations, while V_i and R_i were determined using the sample with the 100 average values[76] for the close equals groups; H_i was finally determined by default. The plotting of the relevant 100 mean values for each of the sets can be observed in the last three graphs on the annex[77].

The final results can be summarized as follows, in terms of the computed absolute values and the corresponding percentage values:

[74] The opposite effect could happen, i.e. direct taxes acting as contributing to diminish global HI, when applied to the XPB distribution, but this does not seem so probable as the inverse.

[75] See sections 3.2 b) i) and 4 i) of this work for further details on the use of the Gini in the measurement of HI and the corresponding formulae and notation.

[76] Sorted by X for sets 1 and 3, and by XPB for set 2.

[77] These show X-t(X) against X, in our earlier notation; they all provide evidence of progression.

	RE_i	V_i	H_i	R_i
i = 1	0.08003	0.0906	0.01017	0.0004
	100%	*113.2075%*	*12.7077%*	*0.4998%*
i = 2	0.04994	0.0541	0.00414	0.00002
	100%	*108.3299%*	*8.2899%*	*0.04%*
i = 3	0.12997	0.1499	0.01683	0.0013
	100%	*115.3343%*	*14.3341%*	*1.0002%*

TABLE: *Israeli benefit (i=1), direct tax (i=2) and net tax (i=3) systems: redistributive effect, vertical and horizontal inequity.*

The analysis of these results leads to a confirmation of our *a priori* intuitions: in respect of Israel, both the benefit and the direct tax systems are progressive (with positive RE), i.e. inequality-reducing and welfare-improving, either when considered as acting "separately" (sets 1 and 2), or jointly (set 3). Moreover, the magnitude of HI is quite relevant, induced by the action of each of the systems (sets 1 and 2), and being reinforced when we consider the joint action of the net system (set 3).

These conclusions, however, should be relativised since, on the one hand, a deeper mathematical study on the possibilities of comparing and relating the above-presented figures is needed, and, on the other, the statistical inference to the global Israeli population is of questionable validity. Further future research on these two topics would, therefore, be useful, as, indeed, would a comparison of the intermediate results with those coming from the alternative sequencing of direct taxes and benefits.

REFERENCES:

Aronson, R., Johnson, P., and Lambert, P. J. (1994) "Redistributive effect and unequal income tax treatment", Economic Journal, vol.104, p.262-70.

Aronson, R. and Lambert, P. J. (1994) "Decomposing the Gini Coefficient to Reveal Vertical, Horizontal and Reranking Effects of Income Taxation", National Tax Journal, vol.47, p.273-94.

Atkinson, A. B. (1970) "On the measurement of inequality", Journal of Economic Theory, vol.2, p.244-263.

Atkinson, A. B. (1973) "How progressive should income-tax be?", in M. Parkin (ed.), "Essays on Modern Economics", Longman, Ch. 16.

Atkinson, A. B. (1980) "Horizontal Equity and the Distribution of the Tax Burden", in H. J. Aaron and M. J. Boskin (eds.), "The Economics of Taxation", Brookings Institution, Washington D. C., p.3-18.

Atkinson, A. B. (1983) "The Economics of Inequality", Oxford University Press, Chs. 1-3.

Barr, N. (1993) "The Economics of the Welfare State", 2nd ed., Oxford University Press, Ch.6.

Barr, N. and Whynes, D. (eds.) (1993) "Current Issues in the Economics of Welfare", Macmillan, Ch.6.

Borooah, V. K. and McKee, P. M. (1994) "Intra-household income transfers and implications for poverty and inequality in the UK", in J. Creedy (ed.), "Taxation, Poverty and Income Distribution", Cambridge University Press, p.69-86.

Brennan, G. (1971) "Horizontal Equity: an Extension of an Extension", Public Finance, vol.26, p.472-87.

Culyer, A. J. (1991) "The Political Economy of Social policy", Gregg Revivals, Ch. 5.

Duclos, J. and Lambert, P. J. (1997) "A Normative Approach to Measuring Classical Horizontal Inequity", Discussion Paper in Economics No. 97/3, University of York, England.

Ebert, U. (1997) "Social welfare when needs differ: an axiomatic approach", Economica, vol.64, p.233-244.

Feldstein, M. (1976) "On the theory of tax reform", Journal of Public Economics", vol.6, p.77-104.

Fellman, J. (1976) "The effect of transformations on Lorenz curves", Econometrica, vol.44, p.823-4.

Fisher, F. M. (1987) "Household Equivalence scales and Interpersonal Comparisons", Review of Economic Studies, vol.54, p.519-524.

Jakobsson, U. (1976) "On the measurement of the degree of progression", Journal of Public Economics, vol.5, p.161-168.

Johnson, S. B. and Meyer, T. (1962) "An Extension of Sidgwick´s Equity Principle", Quarterly Journal of Economics, vol.76, p.454-63.

Kakwani, N. C. (1976) "Measurement of Tax Progressivity: an International Comparison", The Economic Journal, vol.87, p.71-80.

Kakwani, N. C. (1984) "On the Measurement of Tax Progressivity and Redistributive Effects of Taxes with Applications to Horizontal and Vertical Equity", Advances in Econometrics, vol.3, p149-68.

Kakwani, N. C. (1986) "Analysing redistribution policies. A study using Australian data", Cambridge University Press, Chs. 2 to 5, 8.

Kondor, Y. (1975) "Value judgements implied by the use of various measures of income inequality", Review of Income and Wealth, vol.21, p.309-21.

Lambert, P. J. (1985) "Social Welfare and the Gini coefficient revisited", Mathematical Social Sciences, vol.9, p.19-26.

Lambert, P. J. (1993) "The Distribution and Redistribution of Income. A Mathematical Analysis", 2nd ed., Manchester University Press.

Lambert, P. J. (1995) "On the Measurement of Horizontal Inequity", Working Paper No. WP/95/135, International Monetary Fund.

Lambert, P. J. and Aronson, R. (1993) "Inequality Decomposition Analysis and the Gini Coefficient Revisited", The Economic Journal, vol.103, p.1221-27.

Lambert, P. J. and Ramos, X. (1997) "Horizontal Inequity and Vertical Redistribution", International Tax and Public Finance, vol.4, p.25-37.

Lambert, P. J. and Ramos, X. (1997b) "Horizontal Inequity and Reranking: a Review and Simulation Study", Research on Economic Inequality, forthcoming.

Lambert, P. J. and Yitzhaki, S. (1995) "Equity, Equality and Welfare", European Economic Review, vol. 39, p.674-682.

Manser, M. E. (1979) "Comparing Households with Different Structures: the Problem of Equity", American Economic Review, vol.69, p.222-26.

Musgrave, R. A. (1990) "Horizontal Equity, Once More", National Tax Journal, vol.43, p.113-122.

Pechman, J. A. (1989) "Tax Reform, the Rich and the Poor", 2nd ed., Havester Wheatsheaf, Ch.3.

Plotnick, R. (1981) "A Measure of Horizontal Inequity", Review of Economics and Statistics, vol.63, p.283-88.

Sen, A. (1973) "On Economic Inequality", Clarendon Press, Oxford.

Sen, A. (1982) "Ethical Measurement of Inequality: some difficulties", in A. Sen, "Choice, Welfare and Measurement", Blackwell.

Shorrocks, A. F. (1983) "Ranking Income Distributions", Economica, vol.50, p.1-17.

Yitzhaki, S. (1983) "On an extension of the Gini index", International Economic Review,

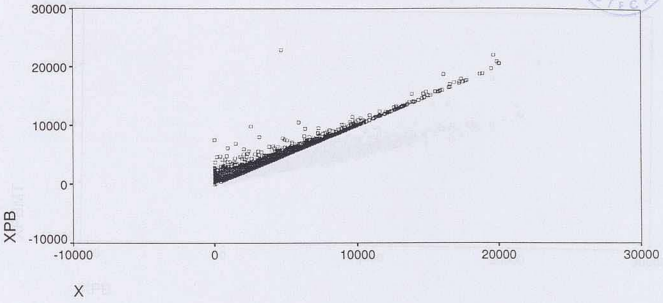
vol.24, p.617-28.

ANNEX

X - Equivalent income

XPB - X after benefits

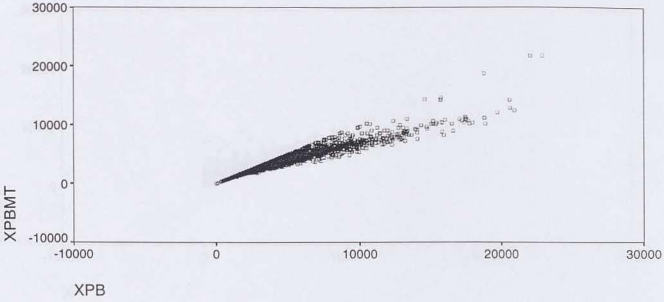
(5200 observations)



XPB - Equivalent income after benefits

XPBMT - XPB after direct taxes

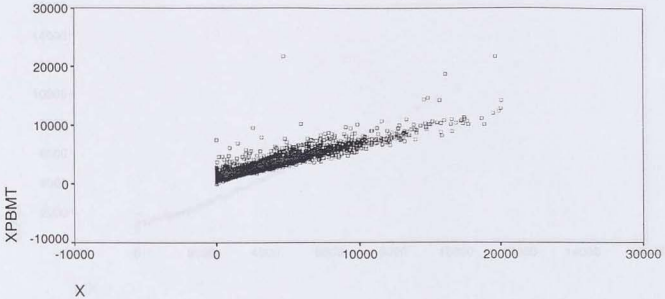
(5200 observations)



X - Equivalent income

XPBMT - X after benefits and direct taxes

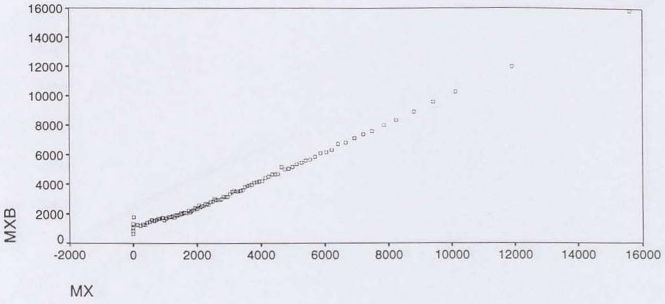
(5200 observations)



MX - Mean equivalent income

MXB - Mean eq. income after benefits

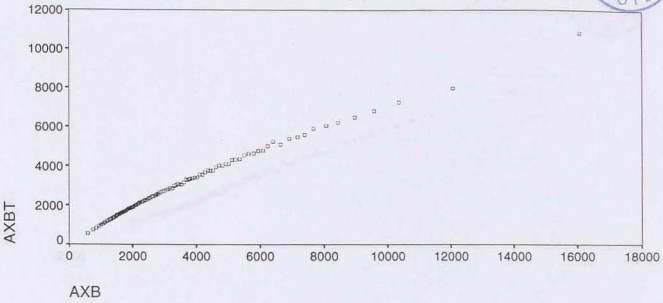
(from 100 subsamples sorted by X)



AXB - Mean equivalent income after benefits

AXBT - AXB after direct taxes

(from 100 subsamples sorted by XPB)



MX - Mean equivalent income

MXBT - MX after benefits and direct taxes

(from 100 subsamples sorted by X)